



# Design of reference populations

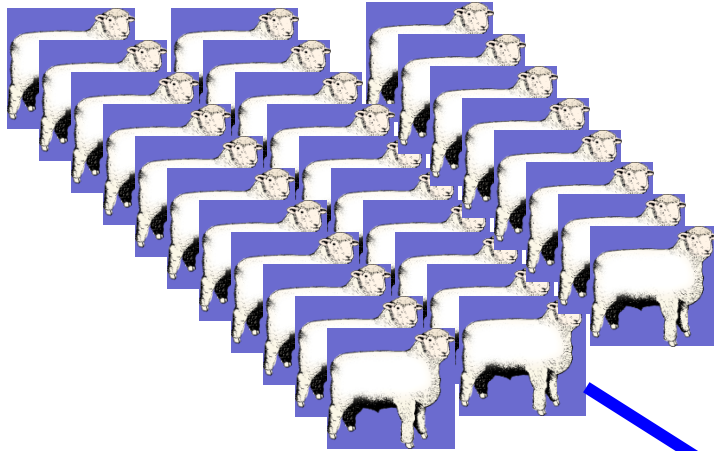
Julius van der Werf



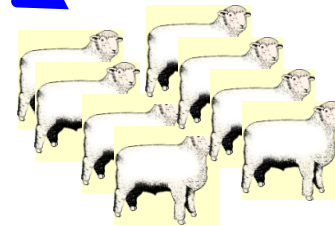
*CRC for Sheep Industry Innovation  
School of Environmental and Rural Science, UNE, Armidale, NSW*



# Genomic Prediction: basic idea



1) Somebody (else) measures lots of sheep, and their DNA  
→ Reference population



2) A breeder tests DNA on young rams

Prediction from DNA → genomic breeding values - GBV

GBV + Current EBV → Improved EBV GEBV

Merit depends on  
trait measurability

# Setting up reference populations

Trait is already measured	Early measurement	Late Measurement
YES	No Need	Use industry data (milk, fertility, late wool)
NO	Create Reference population (slaughter)	Create Reference population

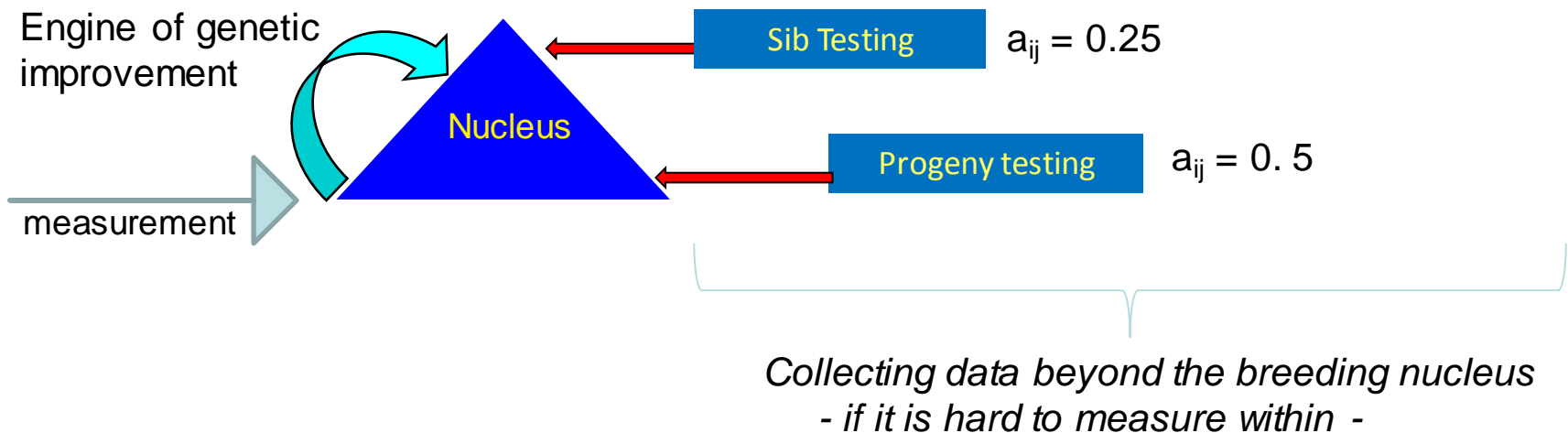
Genomic selection has affected the need for phenotyping !

more...not less

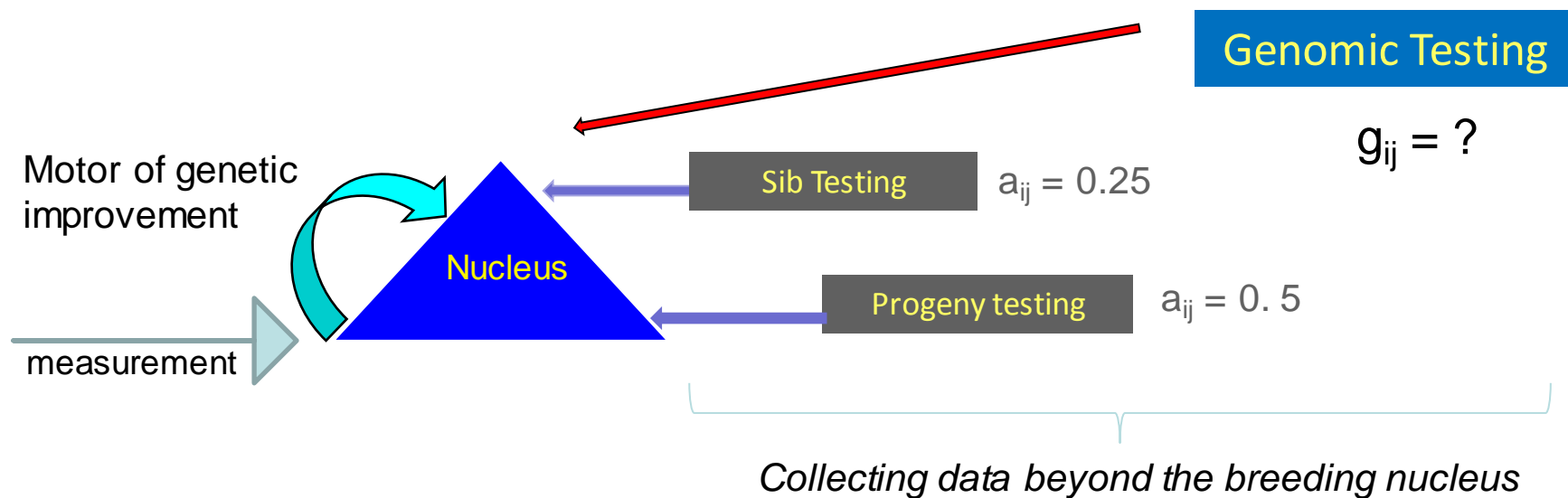
Who pays?

# Design of a reference population

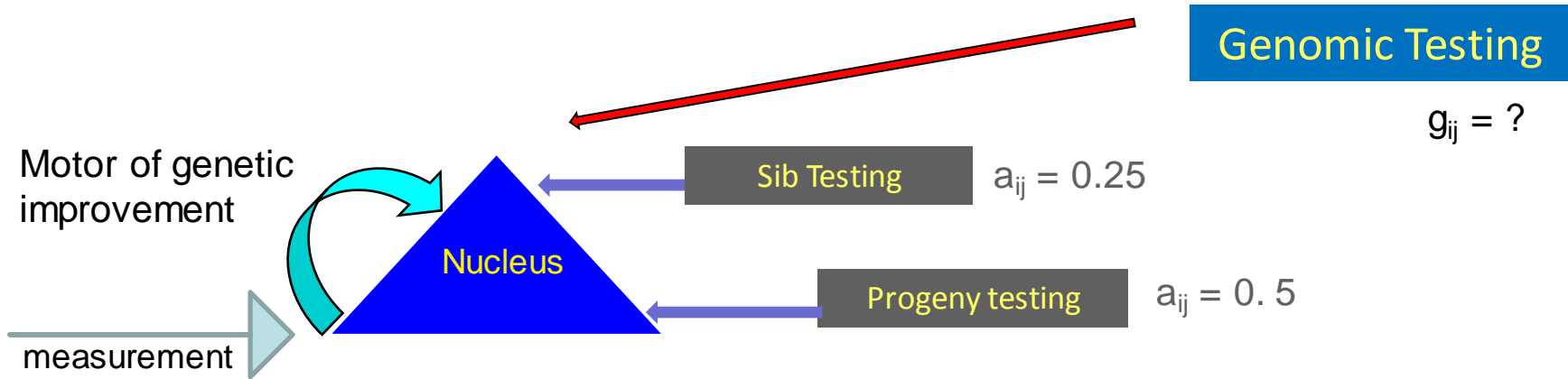
# Investing in information for genetic improvement pre-genomics



# Investing in information for genetic improvement



# Investing in information for genetic improvement



Measure outside nucleus if traits

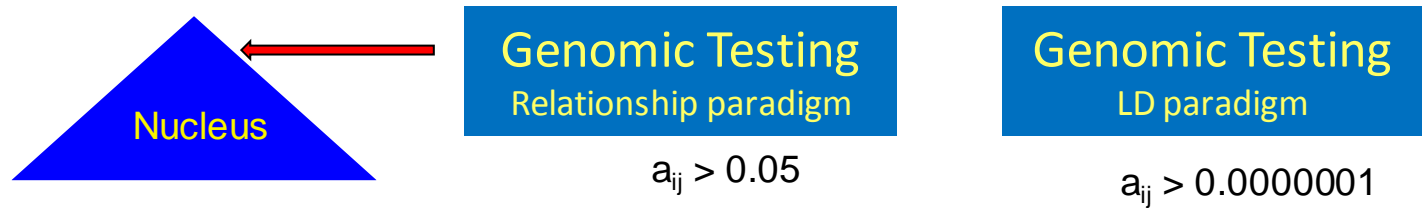
- can not be measured within nucleus
- carcass, eating quality, reproduction

otherwise, reference population can be nucleus

Genomic selection has an advantage over sib or progeny test selection because

1. the information comes earlier
2. can afford to test more distant relatives

# Design of Reference Population



## Relationship paradigm

- Need relatives in reference
- Need to keep reference 'up to date'
- Denser markers maybe of limited benefit
- Accuracy limited by relationships and # of relatives
- Consider to use IBD inference

## LD paradigm

- May achieve prediction across breeds
- Reference population of long lasting benefit
- Accuracy limited by marker density and size of reference
- Requires detectable average effects  
across wide range of genetic background

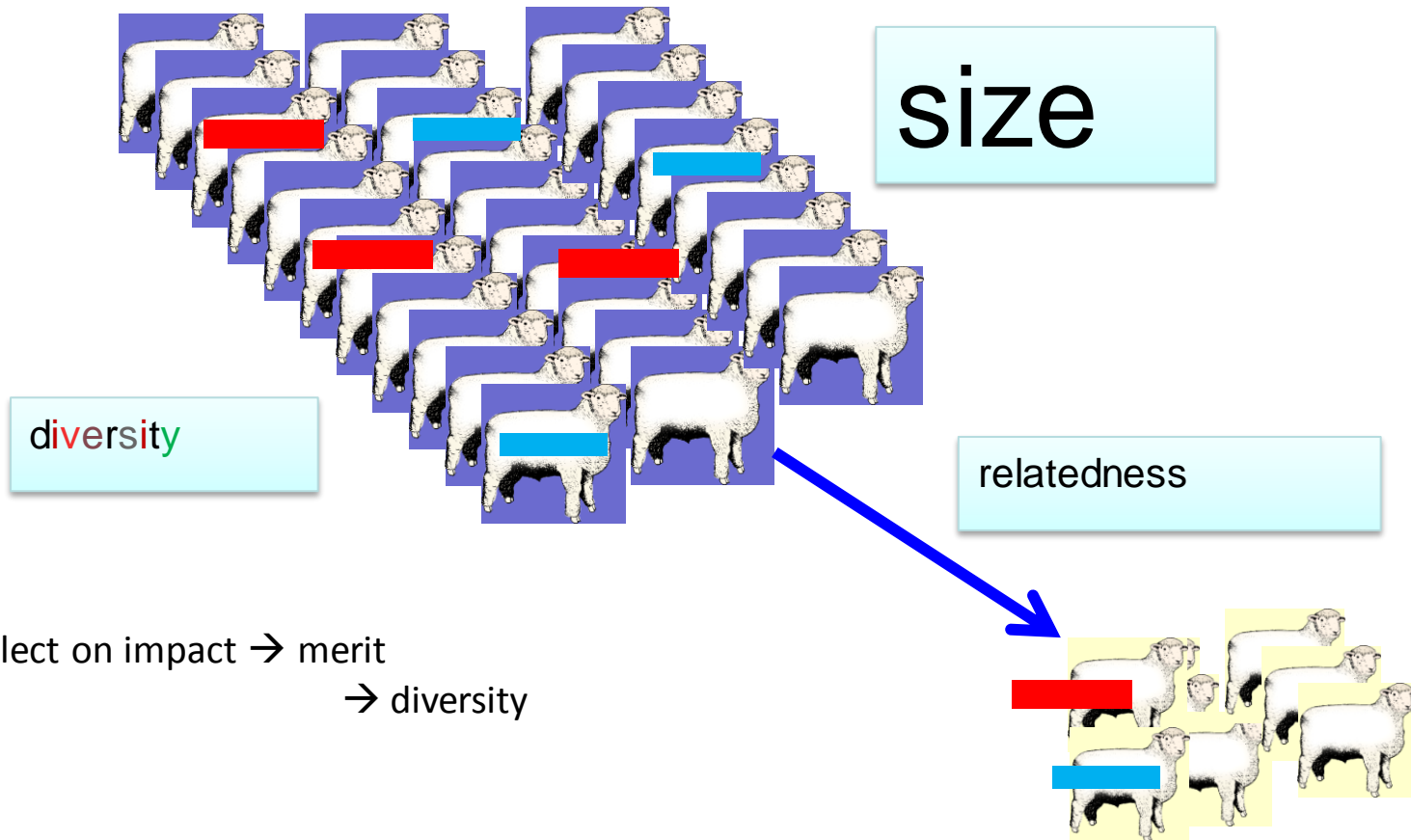


# Summarizing Genomic Prediction

## - What information is used?

- Based on very many small – genomic- relationships
- Does not require ‘direct relatives’ to be tested
- Can be based on distant relatives ‘some generations away’
- .....but the number of small relatives needs to be large (thousands)
- Can not predict across breed

# Design of reference populations

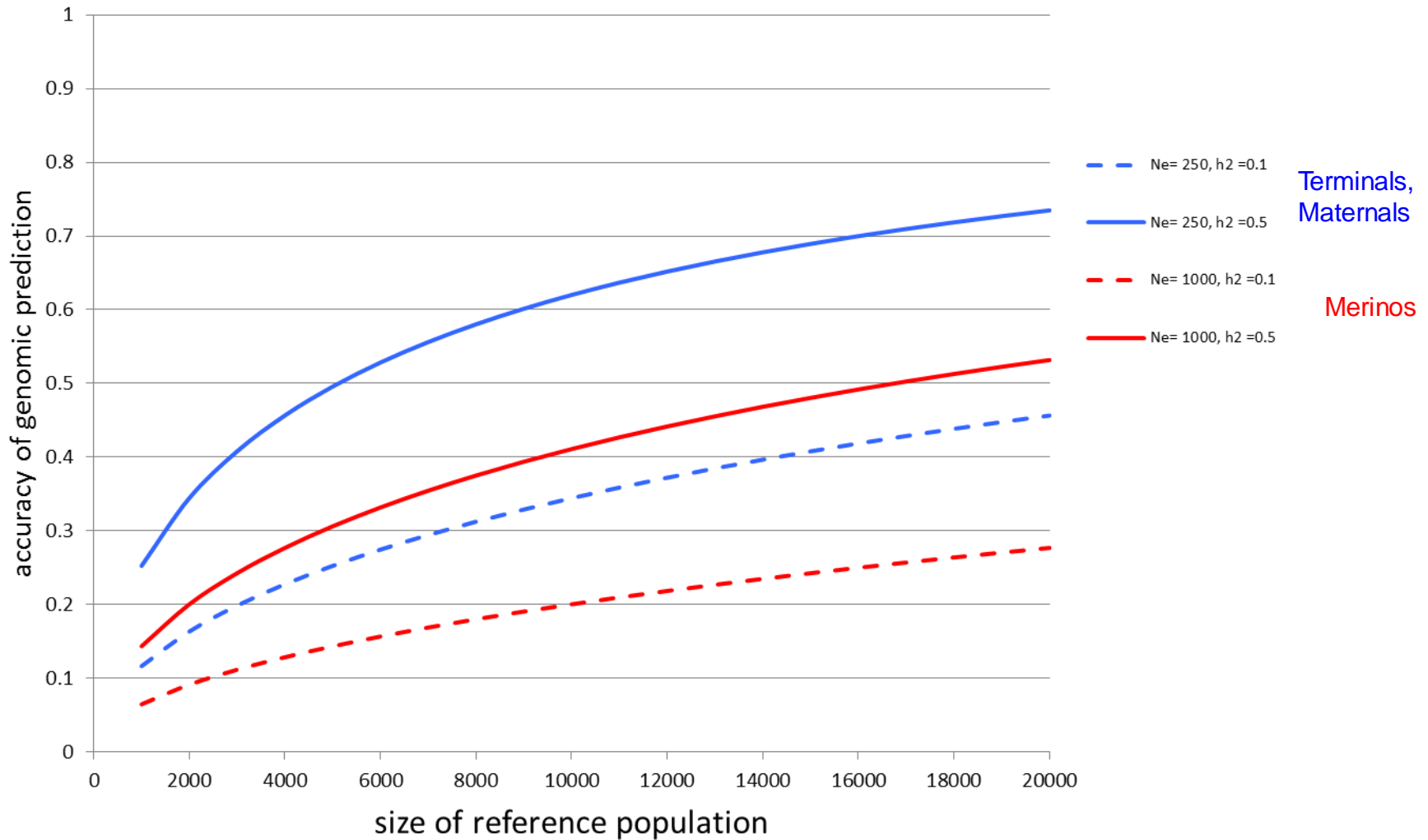


Select on impact → merit  
→ diversity

Multi-breed  
Across breed?  
Longevity of RefPop?

# Accuracy of genomic prediction depending on size of reference population

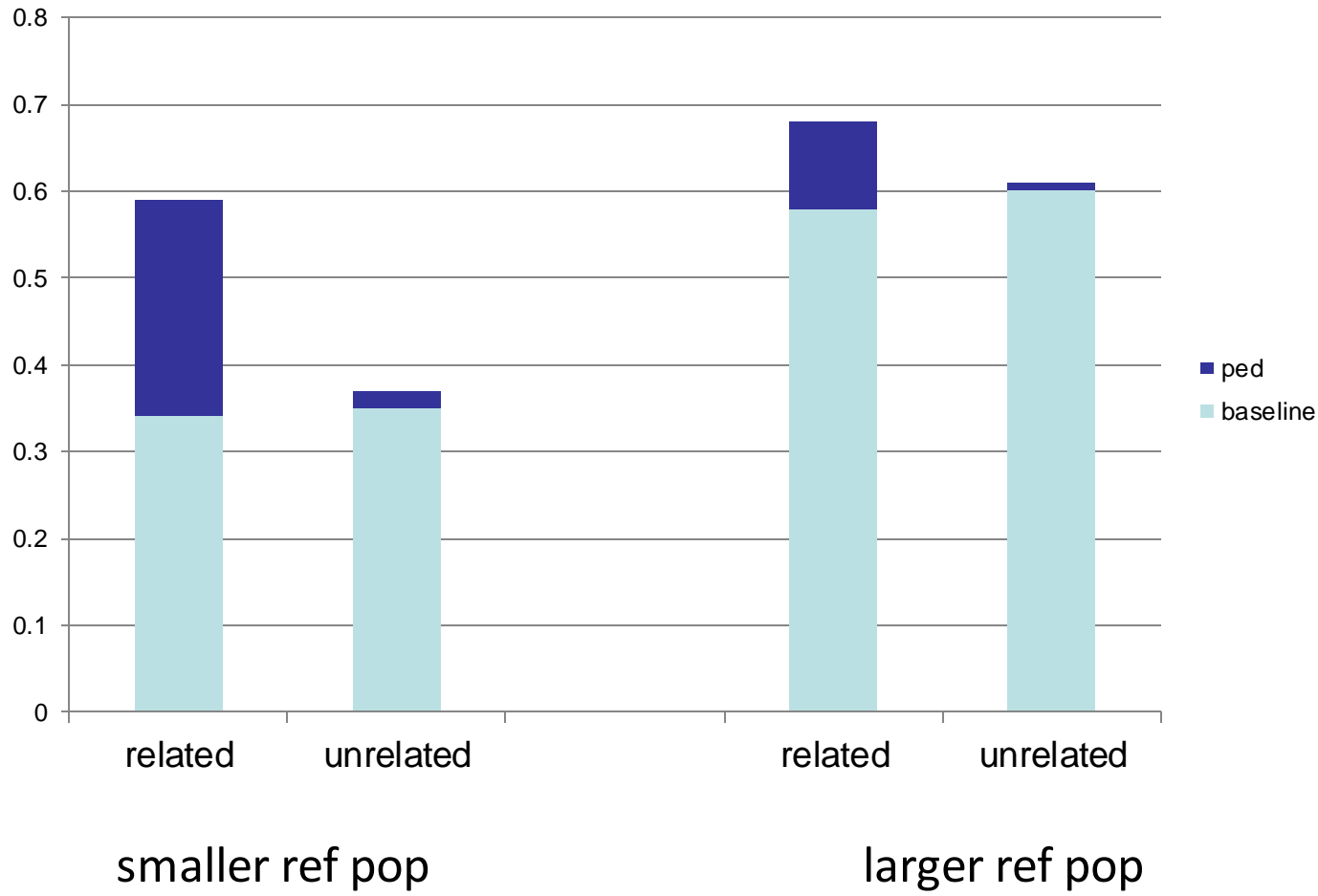
*Goddard 2009*



## design of reference population

- Relatedness between reference population and selection candidates
- Across breeds or lines?
- Number of sires, nr of progeny per sire, which dams?

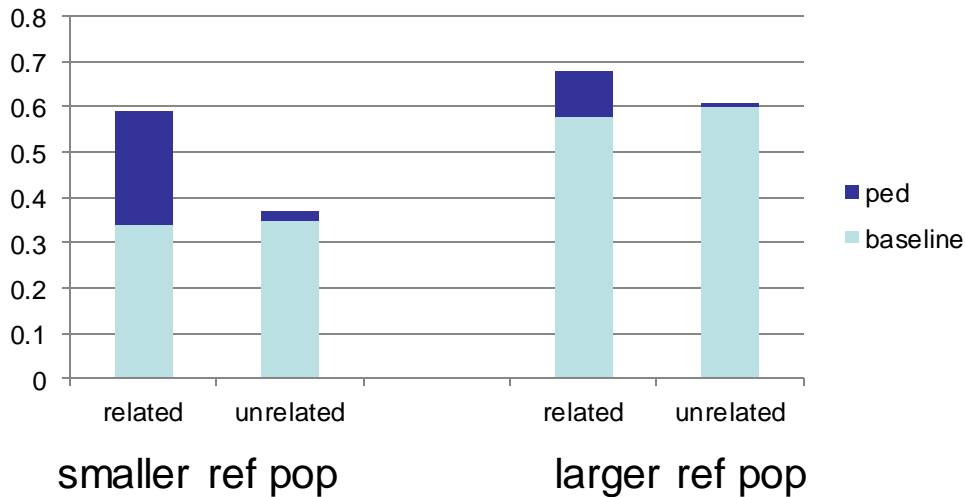
# Relatedness matters more if the reference population is smaller



# Sources of information contributing to GBV accuracy

1. Variation between families
2. Variation within families
3. Markers tracking effects of genome segments/LD  
*Info on 'unrelated'*

BLU P	GBLUP	half life
++	++	1 gen
0	+	1 gen
0	+++	several gen's



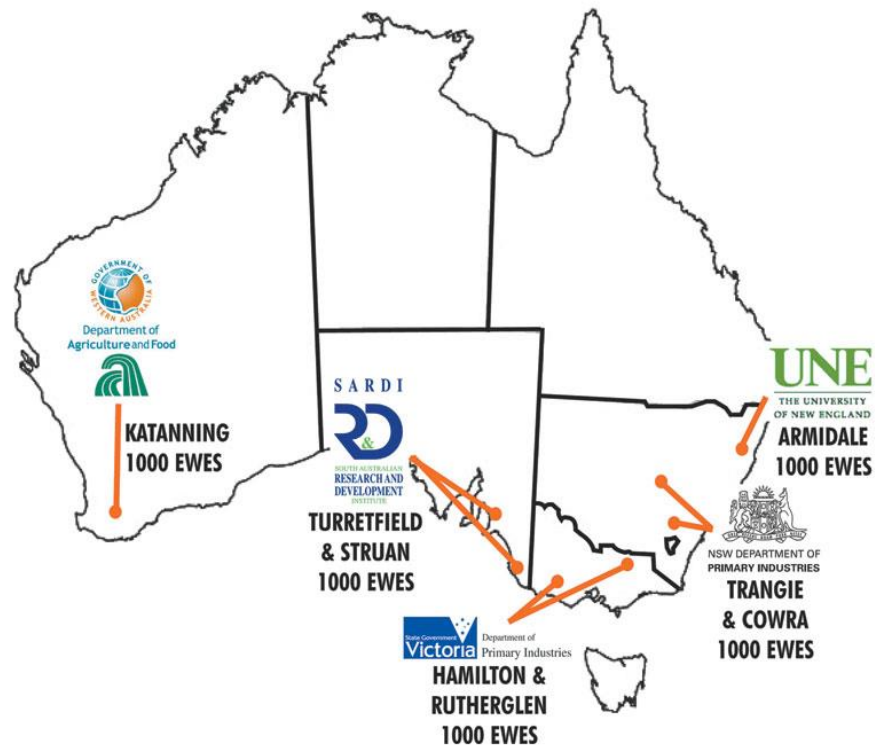
Depending on size of reference population

# Information Nucleus

\* structure \*

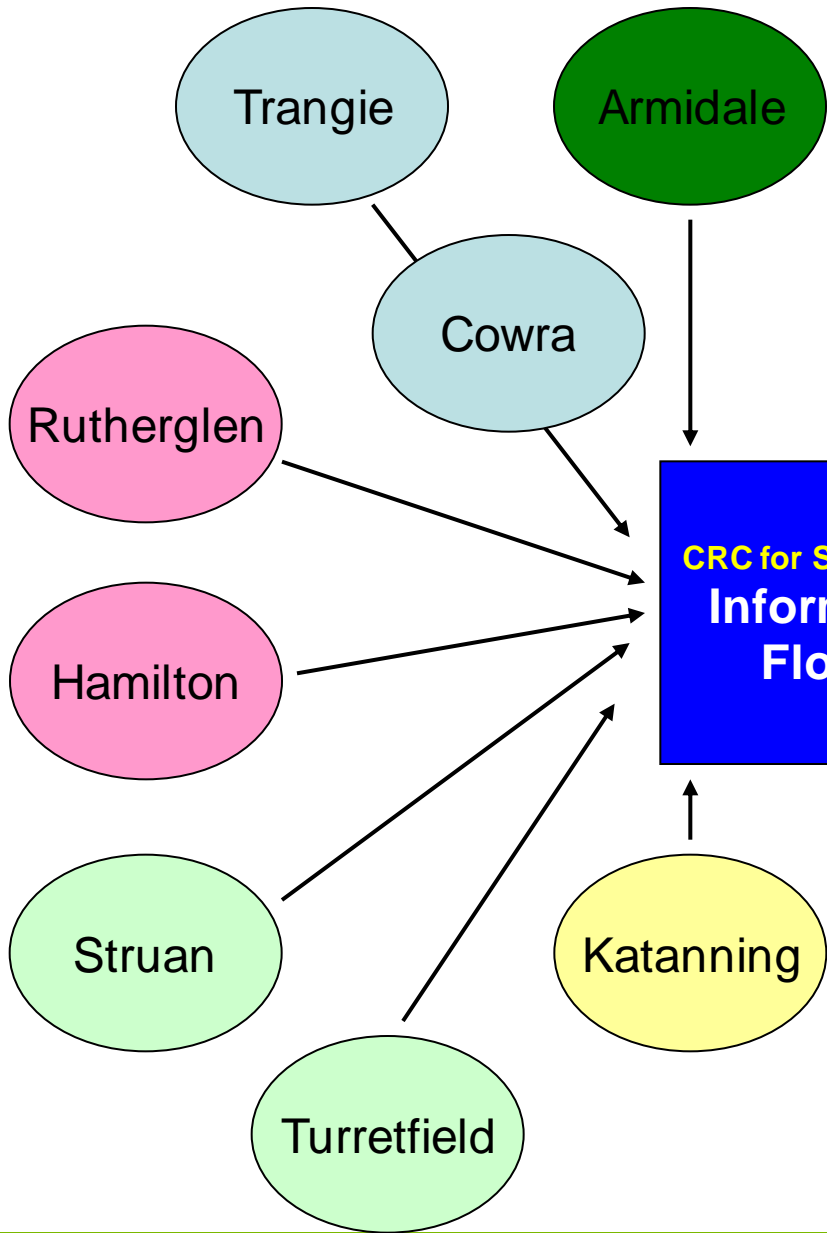


# Information Nucleus Sites





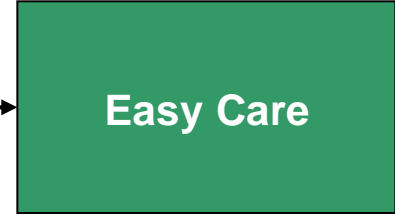
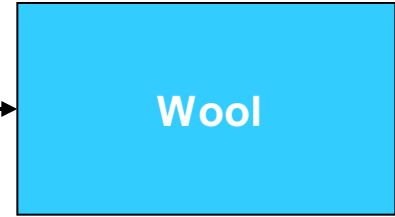
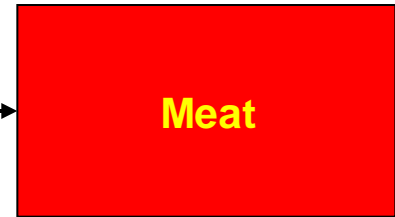
Information Nucleus Flocks



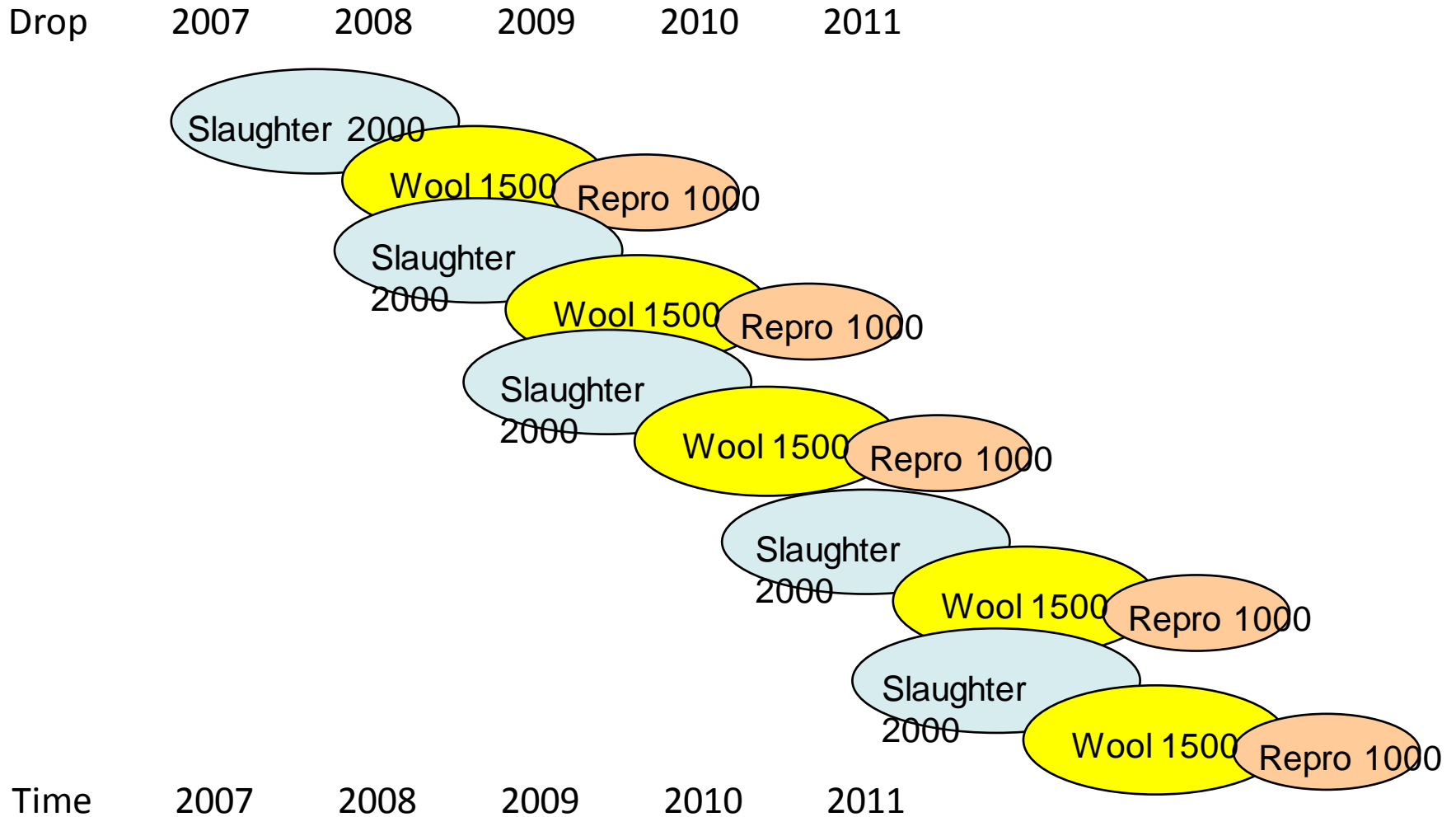
SHEEP GENETICS AUSTRALIA



Sheep CRC Programs



# INF Data Analysis



# How many are needed?

	Breed	Merino	WS,PD	BL	GBV accuracy
	Ne	1000	250	100	
Size Reference pop'n		30,000	10,000	5,000	high
Progeny per year		3750	1250	625	
Predicted benefit in dG		40%	20%	?	
Size Reference pop'n		12,000	4,000	2,000	moderate
Progeny per year		1500	500	250	
Predicted benefit in dG		20%	10%	?	

Assuming the ref pop is refreshed every 8 yrs  
 Numbers are needed per breed!

# INF2 design

breed	rams	prog/ram	Meat	Wool	Repro	Xbred ewes	Merino ewes	Total ewes
WS	72	13	896			448	560	
PD	72	13	896			448	560	
BL	27	30	101		403	202	756	
Mer SF	27	30	202	484	403		1008	
MerMed	27	30	202	484	403		1008	
MerStrong	27	30	202	484	403		1008	
Total genotyped			2499	1452	1613			
Total ewes						1098	4902	6000
Totals	251		2499	1452	1613	1098	4902	

Note that most terminal progeny are crossbreds  
 A crossbred counts for a half (towards each breed)

# Smaller breeds options

(Assuming a small effective population size)

- Measure 250 progeny per year (from ~ 20 sires)  
or
- Measure + genotype 50 progeny tested sires per year
- Need 8x more to get started (or wait 8 years)